AD_____

Award Number: DAMD17-03-C-0053

TITLE: Practical Implementations of Real-Time Heart Rate Variability

PRINCIPAL INVESTIGATOR: Antonio Sastre, Ph.D.

CONTRACTING ORGANIZATION: Midwest Research Institute
Kansas City, MO 64110

REPORT DATE: March 2004

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 2004	Annual (25 Feb 2003 - 24 Feb 2004)

4. TITLE AND SUBTITLE

Practical Implementations of Real-Time Heart Rate Variability

5. FUNDING NUMBERS

DAMD17-03-C-0053

6. AUTHOR(S)

Antonio Sastre, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Midwest Research Institute
Kansas City, MO 64110

E-Mail: asastre@mriresearch.org

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

Original contains color plates: ALL DTIC reproductions will be in black and white

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

Heart rate variability (HRV) is a useful, non-invasive indicator of autonomic nervous system responsiveness that can be used to signal the need for life-saving interventions, but to date it has not been possible to use it in real-time (RT). Because HRV reflects homeostasis in thermoregulation and blood pressure control, it provides a non-invasive "window" into these processes. The overall objectives of the present research program are to identify aspects of HRV that are predictive of adverse consequences in hypo-and hyper-thermia; (2) to implement HRV metrics for quantitation of these aspects; and (3) to develop RT HRV collection software and RT computation of the custom HRV metrics under field-like conditions using modest capability processors. To those overall objectives, the specific aims of the research are to: (1) Acquire temperature and HRV data on human volunteers approaching hypothermia or hyperthermia, to develop custom HRV metrics predictive of adverse outcomes; (2) Develop a software system capable of processing HRV data in RT under field-like conditions and the metrics developed in (1); (3) Validate the software and metrics in RT with a different and larger cohort of volunteers.

14. SUBJECT TERMS

heart rate variability, real-time, hypothermia, hyperthermia, algorithm, R-wave identification, human volunteers,wavelets

15. NUMBER OF PAGES

37

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

INTRODUCTION:

Heart rate variability (HRV) is a useful, non-invasive indicator of autonomic nervous system responsiveness that can be used to signal the need for life-saving interventions, but to date it has not been possible to use it in real-time (RT). Because HRV reflects homeostasis in thermoregulation and blood pressure control, it provides a non-invasive "window" into these processes. The overall objectives of the present research program are to identify those aspects of HRV that are predictive of adverse consequences in hypo- and hyperthermia; (2) to implement metrics or decision rules for detection of these aspects; (3) to determine the rules' false-positive and -negative rates, and (4) to develop RT collection software of HRV in different modalities and RT computation of the rules using modest capability processors. In order to accomplish those overall objectives, the specific aims of the research are to: (1) Acquire temperature and HRV data (using multiple modalities) on human volunteers approaching hypothermia or hyperthermia, to first develop and then validate HRV metrics predictive of adverse outcomes by specificity and sensitivity analyses; (2) Develop a software system capable of processing HRV data in RT and the metrics developed in (1); (3) Validate the software and metrics in RT with a different and larger cohort of volunteers.

BODY:

The objectives and goals of this project are an interactive set of activities:

- Development of novel and challenging software for the RT computation of HRV under noisy, field-like conditions,

- Performance of human studies that will collect data on the responses of the cardiovascular system, and of HRV in particular, to environmental temperature extremes, and

- Development of custom HRV metrics, that will use the custom RT-HRV software, to help predict the onset of hypo- or hyper-thermia.

In order to accomplish these goals in an orderly fashion the project was divided into specific Tasks, but many of the Tasks interacted and cut across the software/human factors divide. As will be detailed below, unavoidable delays were encountered in being able to initiate the Protocol for collection of new data from human volunteers, and the Contracting Officer's Technical Representative and the Project Officer were kept fully apprised of the delays and their causes.

However, in order to maintain orderly and timely progress, software development was kept at full pace. At all the places where software development called for new data that would have come from the performance of the study with human volunteers, we used whenever possible ECG records that had been previously collected by the PI. The data that was needed was of two kinds, and one type could be replaced. We needed to collect data on HRV responses to temperature extremes, and under noisy, field-like conditions, and insure that we extracted the ECG signal reliably in spite of the noise. There was no pre-existing data on the HRV responses to temperature extremes. But there was considerable data in the PI's digital archives of data with

varying levels of noise, including data collected from volunteers engaged in heavy exercise, which would exceed in the levels of noise expected in the data collection in the human subjects phase of this study. These data were used, and that allowed for the software development to be at, or ahead, of the projected timeline.

The approved Statement of Work (SOW) for this project consisted of the following five tasks:

Task 1: Protocol Development—Human Studies (Study 1 and Study 2)
Task 2: Development of Real-time Acquisition and Processing Software
Task 3: Noise Sensitivity Analyses
Task 4: Performance of Human Studies (Study 1 and Study 2)
Task 5: Final Report and Publications

TASK 1. Protocol Development—Human Studies (Study 1 and Study 2). A detailed protocol was prepared for both Study 1 and Study 2. In this report, we provide a summary of the characteristics of the Protocol that has been developed, and of its tortuous review history with the Surgeon General of the Army's HSRRB. Only issues of scientific importance and/or issues that contributed to significant time delays are mentioned, and numerous but minor administrative issues have been omitted.

As noted in the approved SOW, Study 1 and Study 2 have exactly the same experimental designs in terms of environmental exposure of human volunteers, but the data in the two studies would be used in different ways, as outlined below. The characteristics of the participants will be the same for both studies, selected for relevance to military populations. The original SOW called that we propose two studies within the Protocol, as had been the case in previous studies the PI had conducted on behalf of USAMRMC and approved by the HSRRB. Study 1 called for recruitment to continue until complete data are available for 20 volunteers, and Study 2 called for recruitment until complete data are available for 38 additional volunteers. The goals of the research program calls for development of a number of analytical metrics for description of HRV changes upon approach to hypo- or hyperthermia that take place in real-time. These analytical metrics, to be developed in Study 1, would then be tested in a predictive, hypothesis-testing fashion in Study 2.

Characteristics of the proposed Studies:

Inclusion criteria are:

Age:	between 18 and 25
Body Mass Index:	body mass index (BMI) less than 26.0
Fitness:	engaging in regular exercise (at least 3 times a week)
General Health:	excellent general health with no chronic disease or disability
Diet:	regular diet
Other:	able to speak, read and write English

Variables to be measured during room-temperature baseline periods and "hot" and "cold" altered environmental temperature sessions include:

Rectal temperature via an ultra-thin flexible medical rectal thermistor
Lead II electrocardiogram (ECG) for determination of HRV
Finger-tip pulse oximetry
Ear lobe pulse oximetry

In both the hot and cold air conditions, the volunteers would be asked to alternate between sitting or lying down, or moving around (their choice), and engaging in brief periods of deliberate and investigator triggered-movement. Specifically, the volunteers will be asked to rotate between sitting, walking and performing a variety of natural movements, and strenuous movements, like stomping on the floor, arm stretching and crossing across the chest, etc.

The point is that this is not an exercise study; the purpose of the strenuous movements will be to deliberately introduce movement and muscle-activity (myoelectric) noise into the recordings. That is in order to provide us with ECG records that will have segments that will be heavily contaminated with noise. It is one of the specific, dual goals of our proposal to not only collect the physiologic data under temperature changes, but also to develop software for RT determination of HRV under realistic, field-like conditions. The first approximation we must make to field conditions is that of volunteers moving and inducing noise in the recording apparatus that, in magnitude, can match or exceed the ECG signal itself. It is then the challenge for our software to extract the ECG from that noise.

For the cold air exposure condition, each volunteer will be exposed, in order, first to 30 minutes of continued ambient exposure (21°C [70°F], at 50% relative humidity [RH]) to serve as a baseline. The air temperature will then be changed over a 15 minute period to 5°C (41°F), at 50% RH. The volunteer's responses to exposure to cold will be monitored minute by minute by the staff, with special attention to core temperature. Exposure will be terminated upon: (1) decision by the medical monitor for any reason; (2) an exposure duration to 5°C (41°F) at 50% RH of 75 minutes (i.e., a total of 90 minutes; 15 minutes transition from the baseline to the target temperature and then 75 minutes at 5°C (41°F) at 50% RH); (3) core temperature nearing 36°C (96.8°F); (4) the volunteer wishing to discontinue for any reason (preference or discomfort), whichever comes first.

This protocol incorporates a high safety margin in terms of monitors and triggers, since the trigger rectal temperature, which will be approached, but not attained, is a full degree Centigrade (about 2 degrees Fahrenheit) warmer than the clinical definition of mild hypothermia.

For the hot air exposure condition, each volunteer will be exposed, in order, first to 30 minutes of continued ambient exposure (21°C [70°F], at 50% RH) to serve as a baseline. The air temperature will then be changed over a 15 minute period to 39°C [102°F], at 50% RH. Water and commercially-available sports drinks containing glucose and electrolyte mixtures (but not caffeine or other stimulant ingredients) will be available *ad libitum* throughout the heat exposure, since responses to dehydration are not part of the hypotheses being tested in this study, and dehydration could be deleterious to the volunteers. Volunteers will be encouraged to maintain

themselves hydrated according to their own level of preference and comfort with the beverages of their choice. Just as for the cold exposure condition, the volunteer's responses to exposure to heat will be monitored minute by minute by the staff, with special attention to core temperature. Exposure will be terminated upon: (1) decision by the medical monitor for any reason; (2) a time of exposure to 39°C [102°F] at 50% RH of 75 minutes (i.e., a total of 90 minutes; 15 minutes transition from the baseline to the target temperature and then 75 minutes at 39°C [102°F] at 50% RH); (3) approach to a core temperature of 39.0°C [102°F]; **(4) attainment of a mean (average) heart rate of 120 beats per minute[1] for a period of 30 seconds or more[2]; (5) attainment of a heart rate of 140 beats per minute for any period of time, however brief, as signaled by auditory alarms)**; (6) upon the volunteer wishing to discontinue for any reason (preference or discomfort), whichever comes first.

The bolded sections or words above are sections or words, with associated footnotes, that were added or modified between the original submission to the MRI IRB and the most recent submission to the Surgeon General of the Army's HSRRB, as will be explained below.

This part of the Protocol also incorporates a high safety margin in terms of monitors and triggers. For example, the trigger rectal temperature, which will be approached, but not attained, is a full degree Centigrade (about 2 degrees Fahrenheit) cooler from the clinical definition of the onset of hyperthermia. In addition, as a point of perspective, the environmental temperature and conditions of the hot exposure (39°C [102°F] at 50% RH) are not particularly unusual summer conditions for the recruitment area for this study. For example, using climate data gathered from three reporting stations in the Kansas City area and from two reporting stations in Lawrence, KS (all within the recruitment area of the Study) from the US Weather Service (**www.ncdc.noaa.gov**), we found that KC reported occurrences of temperatures of 100°F or higher 57 times in the past 10-yr period. The Lawrence area (also within the recruitment area of this study) has more limited data in the Archives, and those stations reported temperatures of 100°F or higher 63 times the past 6-yr period.

The first Revision of the Protocol was prepared according to the approved SOW, submitted to the MRI IRB, which approved it May 16, 2003, with corresponding Informed Consent Document, Recruiting Poster and other necessary Human Subjects instruments. All of these materials were then forwarded to the Surgeon General of the Army's HSRRB.

---

[1] The dividing line from a heart rate in normal range to tachycardia is 120 beats/min. Another commonly used stop measure with respect to mean heart rate, that of exceedance of 60% of the age-predicted maximum heart rate, is essentially identical and redundant to the one we are using here due to the age range of our volunteers. That is because the age-predicted maximum heart rate is obtained by subtracting the age of the volunteer from 220; for our volunteers aged 18-25 that would give stop criteria of 117-121 beats/min.

[2] The reason for using a criterion of 30 seconds is that a volunteer may **transiently** elevate his/her mean heart rate above going from a resting, supine position to standing and performing the requested natural movements, like stomping on the floor, stretching, moving and waving their arms, in order to deliberately induce myoelectric noise in the electrocardiographic electrodes. Such a transient elevation, lasting only a few seconds, would not be considered indicative of an adverse event. In contrast, a sustained elevation in mean heart rate **could** signify a potential problem, and it is therefore included as a stop criterion even in the absence of core temperature elevations or any reports of subjective discomfort.

However, preliminary review of the Protocol and related materials by the Human Subjects Protection Scientist at the HSRRB in charge of this project indicated that we would not be able to proceed as originally planned in the approved SOW. The PI was informed that, prior experience with the HSRRB notwithstanding, protocols for Study 1 and Study 2 would need to be separated, and a separate Protocol for Study 2 submitted after Study 1 had been completed and it had been confirmed that no adverse events had occurred in the performance of Study 1.

In addition, there were a large number of requests that were made for additions or modifications to the Protocol before it could be forwarded to the HSRRB. These included: extensive documentation of facilities, power analysis, stop criteria independent of core temperature, and diagnostic aids.

Documentation. Some Protocols from other institutions to the HSRRB had presented environmental chambers with limited or restricted access/egress to the volunteers, and generally poor conditions for them, and the PI was asked to document fully and extensively all of the facilities, with photographs, although MRI's environmental chamber had been used with no adverse events in two studies previously approved by the HSRRB. A nine-page Appendix with 9 color photographs was prepared that provided a pictorial "walk-though" of the facility and its various safety and comfort features for the volunteers. In addition, the body of the Protocol was supplemented with requested screen photographs of the data collection software (which was still under development), as well as of some of the safety features built-in the some of the data collection apparatus.

Power Analysis. A power analysis justifying the number of volunteers was also requested. Appropriate power calculations are difficult to perform for this study. The reason is that the endpoints of interest—HRV—have received virtually non-existent attention in the scientific literature in the context of environmental changes in temperature. Hence, critical quantities for any power calculation, such as **effect size** and the **variance of the effect size**, are difficult to come by. At the time the original proposal was written to the Peer-Reviewed Medical Research Program, the PI had been able to identify via Medline/NLM, for the approach or attainment of hypothermia, only 5 citations, 2 on patients suffering from poikilothermia, 2 on obstetric surgical procedures (neither of these 4 of any relevance to this Protocol), and the only report of any relevance to this protocol is the pivotal study by Fleisher and colleagues (1996). Likewise, for the approach or the attainment of hyperthermia, a PubMed/NLM search at that time identified only 3 citations, 2 on patients suffering from thyroid disorders, and one very brief report from Germany in the bio-electromagnetics literature on responses to heating with infrared A. Those are not relevant to the conditions that will be used in the proposed Protocol.

This is in marked contrast to other endpoints such as mean heart rate, but not HRV, for which many dozens of relevant reports can be readily identified. For these reasons, the Medical Research Program Peer Reviewers noted in their review: "Almost anything significant that they detect for HRV in thermal disturbances will be novel." (p. 5 of the peer review). This situation has not materially changed since the proposal was written in June of 2002.

In order to provide a defensible power calculation, which of necessity will be limited by the reliability with which the effect size and the variance of the effect size can be estimated, the PI

and a biostatistician relied primarily on two sources of information. The first is the aforementioned report by Fleisher et al., which explored in a systematic way changes in HRV in human volunteers to surface and to core cooling. The latter, being attained by infusion of chilled saline, is not relevant to the test conditions of the present protocol, but the surface cooling conditions that they used are somewhat relevant, so that their effect sizes and variances of their effect sizes can be used. In the cooling condition, the target endpoint was "shivering," with physiological endpoints being spectral measures of HRV that included very low frequency (VLF), low frequency (LF), and high frequency (HF).

In addition, Fleisher et al. conducted some trials using heating blankets with a target endpoint of induction of perspiration ("sweating") and the same physiological HRV endpoints as for the cooling conditions, and these were used to obtain tentative power calculations for the approach to hyperthermia condition, although these data were not as robust as data obtained from a study conducted by the PI that used moderate temperature exposure of 35°C (95° F) in the same environmental chamber proposed here, under a Protocol approved by the HSRRB (Log No. A-9728).

In that study the goal was to examine the steady-state responses in a wide variety of psychophysiological tests of volunteers who were taking either pyridostigmine bromide or placebo when tested at either an slightly elevated temperature of 75°F or more elevated temperature of 95°F. While a volunteer will equilibrate to steady-state from ambient temperature of 70°F to 75°F in a few minutes, equilibration to steady-state from ambient temperature of 70°F to 95°F could take over 1 hour. To shorten the period of equilibration, we used a heat overshoot method previously developed and calibrated at MRI. The environmental chamber was initially set at 110°F, the volunteer was then exposed to that temperature for a pre-determined time (based on the volunteer's weight), and then the chamber was lowered to 95°F within one minute. The equilibration time at 110°F under no circumstances exceeded 10 min. This procedure had previously been found to be well-tolerated by volunteers, and was found to be very well tolerated in our study, and there were no adverse events and no adverse reactions associated either with this procedure or with any phase of the study.

The total time in the environmental chamber for the performance of the psychophysiological tests was about one hour, but only 9 minutes of supine ECG and 9 minutes of standing ECG were part of that protocol, and all of those data were at steady state. As a result, no ECG data were collected in that study that would document how the ECG changed as the volunteers acclimated to the temperature change. Thus, those data when the volunteers took placebo can be used as surrogates for the power calculation for the Protocol, but are no way duplicative of the types of data needed for the present hypotheses being tested in the present investigation. The data of the sessions when the volunteers took pyridostigmine bromide were excluded from the power calculation.

Power analysis was conducted using these data using standard methods with a target alpha level of 0.05 and power of 0.80 to determine the number of participants needed in order to detect significant differences in measures of HRV under conditions of both heating and cooling as compared to baseline temperature conditions. Since the temperatures, both in the heat (35°C [95°F] in our previous study vs. our proposed target of 39° C [102°F]), and the cold conditions of

Fleisher et al. differ from the target proposed in the Protocol, the calculations must be considered estimates only.

The data presented by Fleisher et al. was first evaluated via an unadjusted effect size index, and subsequently with an adjusted effect size index, for VLF, LF and HF. This unadjusted index was computed for the comparison between baseline and heating ("sweating") conditions and between baseline and cooling ("shivering") conditions. However, given the repeated measures design of the planned study in the Protocol which data are collected from each person serves as his/her own control under baseline conditions as well as under exposure to heat and exposure to cold conditions, the computed effect size was adjusted using a correlation coefficient (r = .694) reflecting the expected correlation between multiple measurements for HRV measures. The overall correlation coefficient (r = .694) was obtained by the conventional method of first computing an individual correlation coefficients (r) for each of the measures (LF, HF, Total Power) for which we had data, computing the *z-transform* of each of the individual r-coefficients, averaging the *z*-transformed r-values, and inverse-transforming the average back to an overall correlation coefficient. This adjustment coefficient was empirically determined using data from previous research performed by the PI and colleagues in which HRV data were collected from healthy young adults under normal and 35° C (95° F) temperature conditions.

The use of this overall correlation coefficient is both justified and necessary in this power calculation, since, as explained in the proposal, and reflected in the approved SOW, it is simply not known which characteristics or metrics of HRV will be sensitive to approaches to hypo- or hyper-thermia.

The adjusted effect size index for the three HRV measures were (a) 1.013, VLF, baseline versus heating; (b) 2.025, LF, baseline versus cooling; and (c) .776, HF, baseline versus cooling. Using the computed effect sizes, it was determined that a sample of 20 participants would be adequate to detect HRV differences under heating and cooling conditions with a power of .80, alpha level = .05, two-tailed.

This power analysis will hopefully serve to satisfy the concerns of the HSRRB, with the best available surrogate data, but it was an extremely time-consuming process to exhaust the peer-reviewed literature for possibly relevant data, and for the actual analyses.

Additional Stop Criteria. In addition to the power analysis, it was requested the PI and the Medical Monitor provide stop criteria for the other than core temperature, (i.e., to be used in addition to core temperature), to add some on-line diagnostic or screening criteria for ascertaining the well-being of the volunteers, and other safety criteria. This was requested even though the stop-exposure core temperatures in the Protocol do not overlap any clinical condition or recognized hazardous condition. For the "cold" exposure, the trigger rectal temperature, which will be approached, but not attained, is a full degree Centigrade (about 2 degrees Fahrenheit) warmer than the clinical definition of mild hypothermia. For the "hot" exposure, the trigger rectal temperature, which will be approached, but not attained, is a full degree Centigrade (about 2 degrees Fahrenheit) cooler from the clinical definition of the onset of hyperthermia.

In order to attempt to satisfy this added request, the PI first conducted exhaustive literature searches through PubMed. In addition, he consulted the Chief Resident in Emergency Medicine at St. Luke's Hospital in Kansas City, MO. St. Luke's Hospital in Kansas City has a Level 3 Trauma Care unit located within a less than 5-minute drive from MRI, and is the Hospital that would handle any emergencies that would ever arise in the area around MRI; in addition, the Medical Monitor has full admitting privileges at that facility. Finally, the Medical Monitor requested the assistance on this issue, at no cost to the project, of Sarah A. Nunneley, Ph.D., who has over a dozen peer-reviewed publications and countless other DoD and NASA reports on her lifetime of work at Brooks Air Force Base on human physiological responses to environmental extremes, especially heat.

The consensus of these investigations was that, for the "cold" exposures, there were no objective signs or symptoms that could be linked to the narrow range of temperatures that would be tested. For the "hot" exposures the situation was even more complicated, since these core temperatures, whether measured rectally or inferred from less accurate sublingual or axillary measurements, are routinely attained or exceeded during routine exercise in non-extreme environments. Therefore, it was impossible to find in the peer-reviewed literature anticipated signs or symptoms, in addition to core temperature, that could be used as triggers for removal of the volunteers from the exposure chamber. Nonetheless, in order to provide an added level of protection against idiosyncratic reactions to heat exposure, the Medical Monitor recommended insertion of two stopping criteria based on mean heart rate (120 beats/min for 30 sec or 140 beats/min at any time), as noted in the bolded sections of the description above.

Diagnostic Aids. It was requested that either the volunteers be medically screened prior to participation, or that some diagnostic aids be added to the Protocol to detect any physiological distress that may be caused by the "stresses" of the environmental temperature exposures. That was even though the volunteer population was highly restrictive and biased towards the very fit and healthy: ages 18 to 25, with a BMI less 26.0, engaging in regular exercise at least 3 times a week and in excellent general health with no chronic disease or disability.

The PI performed a cost analysis, since the fees for a pre-screening medical exam had not been anticipated nor budgeted for in the proposal nor authorized in the SOW. In discussion with the Medical Monitor and with the Human Subjects Protection Scientist, is was suggested that a compromise that may be acceptable to the HSRRB was to incorporate into the protocol a Diagnosing ECG (12-lead) using an FDA-approved unit with the Macfarlane / Glasgow Royal Infirmary Algorithm.

Macfarlane and his colleagues at Glasgow have developed for several decades detailed algorithms that examine a 12-lead ECG, and using normative data for age, gender, ethnic background, and a whole host of clinical conditions, can give conservative diagnoses and summaries. Their algorithm has been incrementally validated over decades in Europe, Asia, and America with several hundred thousand patients, and several implementations, including the Burdick Diagnosing Elite II, have FDA approval for full clinical use in the US.

Cost-wise, this was a much more efficient and reasonable way to comply with the request. In the protocol, we will periodically activate the Diagnosing 12-lead ECG Burdick Elite II unit

with the Macfarlane / Glasgow Royal Infirmary Algorithm. This unit requires, for avoidance of false positives, that the volunteer be lying down in a couch and remain fairly quiet and movement-free for the period of 12-lead diagnosing data collection. During the 75 minutes of the active hot or cold environmental exposures, we will ask the volunteer to lie down every 15 minutes and we will activate the Diagnosing ECG unit; this will provide five diagnosing tests during the test exposure.

The output of the Diagnosing 12-lead ECG Burdick Elite II unit automatically switches from displaying electrocardiographic Leads II, I, II, aVR, aVL, aVF, V1, V2, V3, V4, V5, and V6. It then shows the summary for the ventricular rate, P and QRS durations, PR, QT, and QTc intervals, and P, QRS and T axes. With this plus other information collected but not displayed, age and ethnic information that was given as input to the algorithms, the Macfarlane / Glasgow Royal Infirmary Algorithm provides its diagnose(s)/summaries with qualifiers with respect to degree of confidence from the data obtained.

The revised protocol that incorporated all of the requested changes was re-submitted to the MRI IRB, where it was again approved and given an "Minimal Risk" designation. The materials have been forwarded to the HSRRB, where at the time of preparation of this Annual Report, they are pending review.

TASK 2: Development of Real-time Acquisition and Processing Software. In this Task, as noted in the approved SOW, algorithms and an implementation strategy would be developed for real-time acquisition of ECG and pulse oximetry data and peak identification. In other words, it is routine to digitize ECG data, and to subsequently identify the peaks of the R-waves, and hence the times of occurrence of the R-waves, off-line. The off-line process routinely involves a step of expert operator review of the software-identified R-wave peaks for peaks that the software may have mis-identified due to noise, or ectopic beats, etc. It is from the R-R intervals that all HRV computations take place.

Challenges: The first real challenge in porting HRV to real-time (RT) is to do the identification of the R-waves in RT, to do it accurately, and to do it eventually with minimal to no human operator editing. The second, and most formidable challenge, is to do it not just for very clean signals, such as can be recorded under laboratory conditions, with very large signal-to-noise ratios (S/N). Rather, one of the goals of the approved SOW was to extend the available technologies of digital filters that the PI had used and proposed to handle ECG signals that were heavily contaminated with myoelectric noise, as is routinely the case when the volunteer is moving or exercising. That is clearly of real-world military relevance. The first step was to apply these concepts and ideas to the ECG, and later to apply them to the simpler pulse oximetry signal.

Finally, in the SOW, real-time computation of HRV metrics will be first based on existing, conventional HRV metrics, with subsequent real-time computation of custom metrics to await until Task 4 when data from volunteers exposed to temperature environmental extremes became available.

Considerable progress has been made on Task 2, with some unexpected successes that exceeded the original expectations. The PI made slight modifications in the original approach to capitalize on these successes.

One of the modifications was not separate Task 2 and Task 3 rigidly (see below), since in fact the SOW called for them to run simultaneously, and efficiencies in time and effort could be accomplished by using insights obtained in the conduct of one Task immediately on the performance of the other. One concrete example of where this approach was applied was in the use of the tool of fractlet filters (discussed in detail below). Logically, these analytical tools can be considered either as a method to implement a more successful RT algorithm (Task 2), or as a way to improve sensitivity to noise (Task 3). The PI decided to apply the fully apply the fractlets, once developed, to both tasks, since they proved useful in both contexts.

The overall strategy that was presented in the original proposal and appeared in the approved SOW for Task 2 was to improve on the Pan-Tompkins R-wave detection algorithm by the use of wavelet-based prefilters. The approach proposed involved:

- Improvements in custom software, including:
    - refinements in the Pan-Hamilton-Tompkins R-wave detection algorithm
    - blending with different pre-filters optimized for performance with the Pan-Hamilton-Tompkins algorithm
- Investigation of alternative R-wave detectors to compare performance in the presence of severe noise.
- Validation of a wavelet library against the commercially-available S-Plus, to be used in our custom software.
- Implementation of a three-stage filtering testbed to allow visual experimentation with filters and peak detectors.
- Development and use of one set of very carefully-scored (aka "Gold Standard") existing ECG records to first test new algorithms
    - This first, test "Gold Standard" set would also be used in Task 3 to assist in the determination and computation of noise sensitivity metrics
- Use of a second set of very carefully-scored (aka "Gold Standard") existing ECG records to validate the new algorithms

Wavelet background: A wavelet transform is a remarkably useful generalization of the Fourier transform. In a wavelet transform, the signal to be examined is repeatedly decomposed by means of a pyramidal decimating algorithm using one of many available, recursively-defined

analyzing wavelets with discrete time-support. At each level of decomposition, half the number of points in the time series are used, resulting in a number of levels of "detail." The levels of detail go from highest detail (i.e., higher frequency content) to lower detail (lower frequency content), each with a factor of two fewer points than the previous level, until a "remainder" is obtained. This is equivalent to passing the signal through a bank of sharply tuned frequency filters. The advantages of the digital wavelet transform (DWT), which account for its great popularity in a number of applications, include: (1) computational efficiency (as fast or faster than the fast digital Fourier transform), (2) the availability of a large number of different wavelet "families" with different properties (e.g., Daublets, Coifflets, Symmlets, bi-orthogonal spline wavelets, etc), (3) the ability to tailor the "width" (or support) in time for a given wavelet family, allowing for multiscale analysis, and (4) the localization of frequency components to specific points in time (i.e., inherent suitability for time-frequency analysis).

In the original proposal the PI presented preliminary data from his laboratory that illustrated an ECG record collected from a volunteer exercising in an elliptical cross-trainer, which generated a large amount of myoelectric noise on the ECG record. As was shown in the proposal, an optimized Pan-Hamilton-Tompkins algorithm failed to adequately identify and locate R-wave peaks. Likewise, the commercial (proprietary) template-matching algorithm present in the Rozinn Holter for Windows® software also failed to identify the R-waves for meaningful HRV analysis in that segment. In contrast, the R-waves in that record were readily identified after first being denoised by a wavelet transform.

In Task 2, the PI worked closely with a subcontractor, Mr. Richard Ulrich (Ulrich Associates, Inc.), to develop custom software for the project, and with the consultant to the project, Prof. Katherine Davis (Math Dept, U. of Texas, Austin), for mathematical and analytical support and solution of problems that was incorporated in the software. During the period of performance, four major releases of the software, each incorporating new levels of functionality, were developed and tested. Each of these major releases included several revisions that fixed bugs or dealt with presentational issues.

New Developments: In the course of attempting to accomplish the dual challenges of R-wave recognition in RT, and the even bigger challenge of R-wave recognition when the R-wave is embedded in very noisy signals, several new developments were made.

The first was the realization that intrinsic to the Pan-Hamilton-Tompkins algorithm, whether as natively-published by these investigators in the open literature or as previously improved and optimized in the PI's laboratory, it still retained an algorithm-specific filter, whose performance was "tuned" to unprocessed ECG records, not to ECG records that had been pre-filtered by wavelet filters. It was therefore possible, even likely, that we could obtain superior performance by optimizing one or more pre-filtering stages with wavelet or other pre-filters, and using only the Pan-Hamilton-Tompkins R-recognition algorithm, without its pre-filter. Hence, we decided to have that Pan-Hamilton-Tompkins pre-filter first as an option, and then removed, from our custom software.

The second new development is what we believe to be a new analytic discovery, of what appears to be strong practical importance for the success of this project. That is the discovery of

"fractional order wavelets, which we have chosen to call fractlets." We have not found these anywhere in the technical literature. We describe below how we came about recognizing the need for a new tool, which resulted in the fractlets, and provide an example of how we feel they will be of practical importance for this project.

We realized, during the testing and validation of our own wavelet library, that for many of the most common and useful wavelet families, like the Symmlet and Coiflet wavelets, the width of the R-wave for a person, or for a given heart rate for that person—since the width of the QRS complex shortens with exercise and with increasing heart rate—falls "in between" wavelet decomposition levels. At noted above, a wavelet transform decomposes the source signal, in this case the ECG, into many levels of "detail," each level having roughly a different frequency content.

It had been our hope to find a single optimal wavelet family, and a single level, that could be used widely across subjects and conditions, and gave acceptably accurate results for R-wave recognition. For example C6d4, meaning a Coiflet with support length 6, level of detail 4. Or, in the absence of finding one particular level that worked all the time, we could have our software's algorithms adaptively change from C6d3 to C6d4. However, any such adaptive shifts always have the danger of a certain amount of computational overhead, and that, in turn, can lead to problems in an RT environment.

The reason why this is an important issue is that there is some data in the literature, and the results of our own Task 3 have begun to bear it out, is that some HRV metrics can be quite sensitive to even small number of missing or miss-assigned R-waves (Xia et al., 1993; Lippman et al., 1994; Berntson and Stowell, 1998). Therefore, finding a filter or set of filters that minimizes the number of mis-assigned R-waves, especially under real-world, noisy conditions, is very important for the success of this project, and for future implications and work that may result from these endeavors.

During a team meeting with the project's consultant, Prof. Davis and the programmer, Mr. Ulrich, the PI inquired of Prof. Davis if anyone had ever reported fractional-order wavelets, and if not, what the analytical hurdles may be to creating a fractional-order wavelets that would have suitably rigorous mathematical properties that would make them useful for improved ECG filtering and R-wave recognition. The answers were that the concept was completely novel, and that the difficulties would not be insurmountable analytically. They all primarily had to do with accurate handling of phase relationships. Mr. Ulrich in turn reassured the PI that there would be only modest difficulties in terms of practical implementation in efficient and fast coding.

Simply put, a fractlet sharpens recognition of sub-octave components in a signal, performing a linear interpolation of phase-adjusted detail filters within a wavelet class. The key to the deceptively-simple definition of "linear interpolation," where a:

Coiflet level 6 of fractional level of detail 2.5 = (0.5*Coifflet6 d2) + (0.5*Coifflet6 d3)

lies in the all-important phase-adjusted detail filters. Without very careful phase adjustment, the results obtained are nonsensical. We extended the properties of our validated wavelet library to allow for the fast and accurate computation of fractlets of all of the supported wavelet families.

A reason why we feel fractlets are an important development for this project is illustrated in the following series of figures, which are screen captures from the most recent version of our custom software, ECG16a. ECG16a is the first version to have two algorithms for R-wave detection, a version of Pan-Hamilton-Tompkins much further optimized than was available at the beginning of the project, and a new method based on correlation of raw and filtered data. It has separate display of previously-validated R-waves against automatically-determined R-wave positions without human operator intervention. The latter are the first attempts at operator-independent R-wave identification, and a necessary prelude to true RT R-wave detection.

The example we are illustrating of the utility of fractlets is not the only example, just one for illustrative purposes in this report. Once we implemented and validated the fractlets was part of our wavelet library, we have found them very useful, and particularly so in dealing with R-wave identification of "real world" records that are heavily contaminated with myoelectric noise.

The ECG file that is being used in the illustrations was collected using a commercial Rozinn Holter monitor at 180 Hz digitization rate, downloaded from the proprietary Rozinn encoding system via Holter for Windows® software using the "research" (highest quality) settings. The proprietary Rozinn binary format file that intermingles 3 channels is converted to three separate files, each with a conventional binary stream encoding of 2-byte integers by a utility kindly provided by Mr. Snay Marchand of Rozinn Electronics, Inc. Once the data are in conventional binary format our software, as well as other software can read and manipulate it. This particular record is what Rozinn labels as Channel 2 in the utility output, which corresponds in our hook-up system to the standard electrocardiographic Lead II. In this previously-available record, the volunteer had rotated through various exercise machines in a gym, and this segment is characteristic of many that are heavily contaminated with myoelectric noise due to his vigorous skeletal muscle activity.

From baseline records where the QRS complexes were clear at slow heart rates, and also from records where the volunteer had exercised to peak capacity and then stopped all motion so that the Holter could record the ECG, and hence the QRS morphology of the exercise-shortened QRS complex, the PI and staff could identify this volunteer's R-waves accurately by careful visual inspection in most of the record, even in noisy segments. The entire record was first scored by one staff member and re-checked by the PI, a procedure that was followed for all of the "Gold Standard" set.

In ECG16a, there are four traces. The first trace is the raw ECG. The second the output of the first filter, which can be "no filter," as in the present example. The third trace has the output of the second filter, as well as other information. On that third trace, the PI-identified R-waves appear as triangles (open or solid), while R-wave assignments that are made by the ECG16a software without using the information in the triangles appears as vertical solid bars. This gives an immediate, minute-to-minute reflection of places where the recognition algorithm failed or succeeded, in addition to tabular summaries for the entire file that can be invoked via menus.

The fourth trace is the output of the integrator (for the Pan-Hamilton-Tompkins R-wave recognition, as is the case here), or the output of the correlation function (if the correlation R-wave recognition method has been chosen).

Figure 1 shows the results of analysis of the aforementioned ECG record in ECG16a, using no Filter 1, a Symmlet of support 4 and level 2 as the filter, and our modified Pan-Tompkins R-wave recognition algorithm. Although it performs rather well for such a noisy record, it fails at second 32.7. It can be seen that the true R-wave (triangle) is missed, and a larger but narrower noise peak is picked instead. While this may seem insignificant, this entire figure only encompasses 4 seconds, and it is known that errors of this magnitude, against the standard accumulation times of most HRV metrics of 5 minutes, can become quite significant (Xia et al., 1993; Lippman et al., 1994; Berntson and Stowell, 1998). There is also a miss at 32.9, but examination of the raw record indicates that is a severely-contaminated portion, and very difficult judgment call to assign the R-wave by the human expert. At that location, the quantitative error between the program and the human-validated assignment is just a few milliseconds, and will not contribute to contaminate or invalidate quantitatively the HRV metrics as much as the missed assignment at second 32.7.

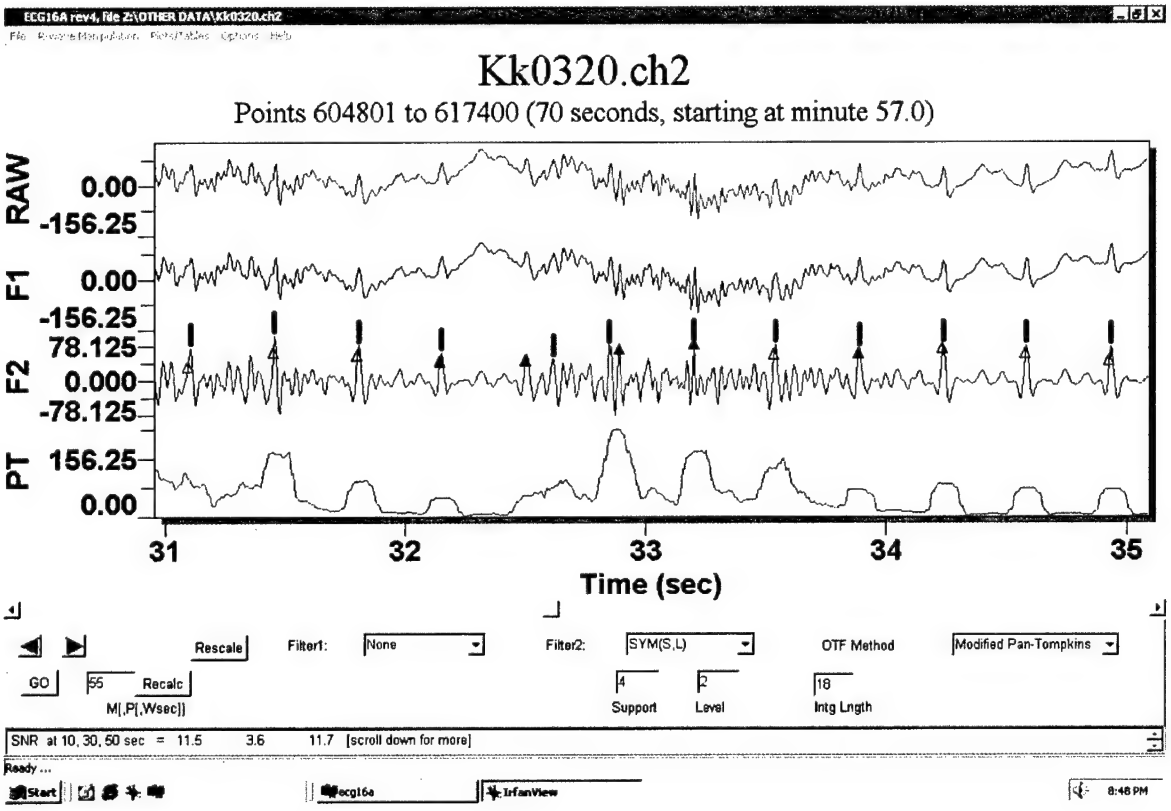

Figure 1. Lead II ECG of volunteer, with a Symmlet of support 4 and level 2 as the filter.

Figure 2 shows the results of analysis of the aforementioned ECG record in ECG16a, using no Filter 1, a Symmlet of support 4 and level 3 as the filter, and our modified Pan-Tompkins

R-wave recognition algorithm. In this case the performance is worse. While the level 3 accurately picks up the R-wave at second 32.7, it completely misses (i.e., it produces a false-negative) at second 33.25. Quantitatively, that introduces an artifactually long, almost double-size R-R interval where the R-wave was missed, and those missed R-waves are known to in turn potentially introduce meaningful artifacts in HRV metrics in the standard accumulation times of 5 minutes (Xia et al., 1993; Lippman et al., 1994; Berntson and Stowell, 1998). As expected, the miss at 32.9 is unchanged.

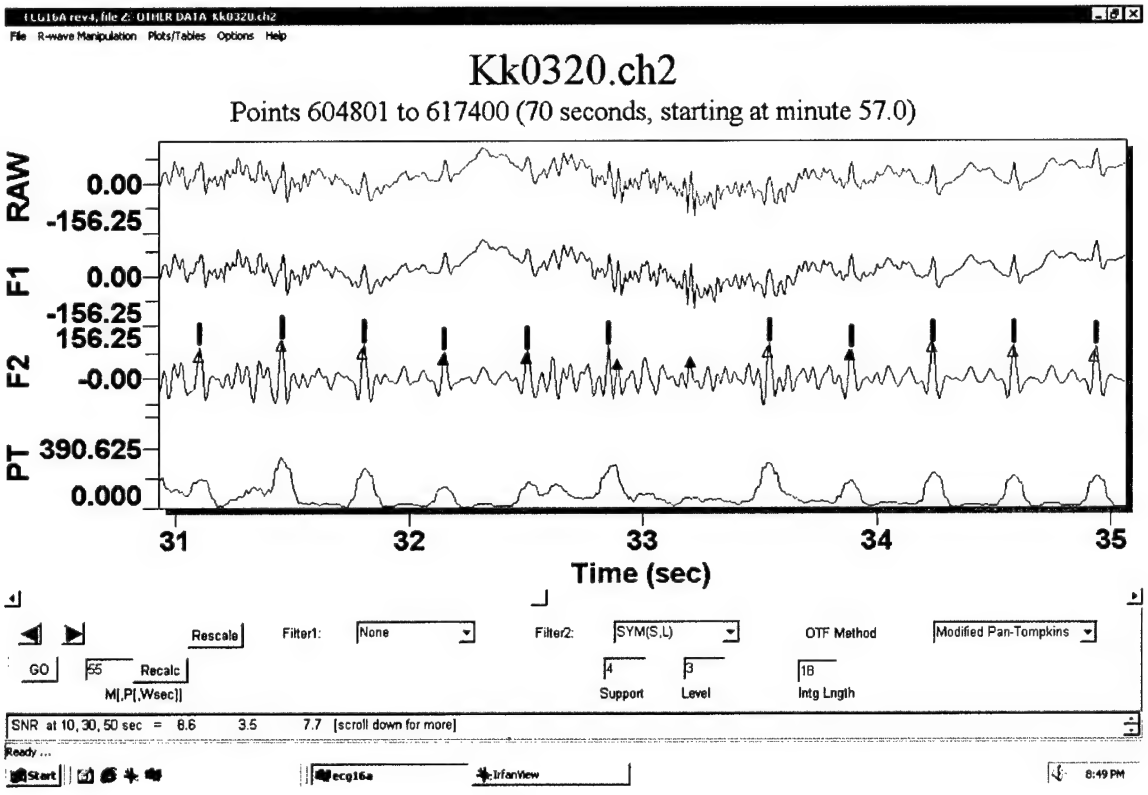

Figure 2. Lead II ECG of volunteer, with a Symmlet of support 4 and level 3 as the filter.

Figure 3 shows the results of analysis of the aforementioned ECG record in ECG16a, using no Filter 1, now using a custom fractlet: a Symmlet of support 4 and level 2.5 as the filter, and our modified Pan-Tompkins R-wave recognition algorithm. All of the R-waves are correctly identified, both at second 32.7 and at second 33.25. As expected, the miss at 32.9 is unchanged. While this may appear as only a minor achievement, when one considers this 4-second record, which is quite representative of ECG records collected when volunteers are moving and exercising, and multiplies the false positive error (of Level 2) or false negative (in the Level 3 detection) against 5-minute accumulation windows, for several hours of HRV determinations, the picture changes.

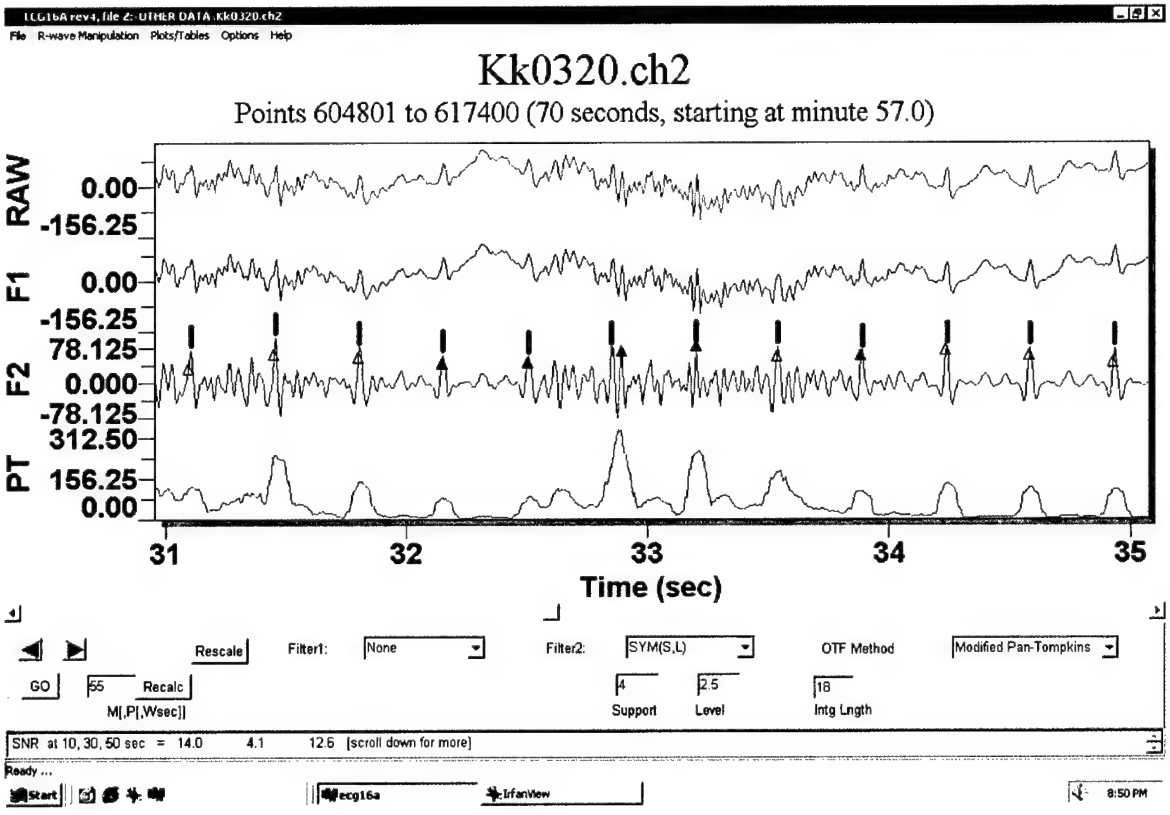

Figure 3. Lead II ECG of volunteer, with the filter being a fractlet: Symmlet of support 4 and level 2.5.

We are confident that fractlet-assisted R-wave detection may lead to computation of much more accurate HRV metrics when data are collected under less-than-ideal conditions. As part of our goals for the next year, we plan to continue to quantify the impact of the missed beats and mis-assigned beats on the HRV metrics' accuracy, and, when environmental extreme exposure data are available, on the accuracy with which body changes can be predicted via changes in HRV metrics.

It is still an open question, an a high-priority item for the next twelve months, to quantify the effects of the remaining mis-assigned or missed beats. Realistically, even a combination of the adaptive modifications to the R-wave detection algorithms we have implemented, which have resulted in significant improvements on their own, coupled with impressive improvements due to the development of fractlet-based filters, still do not lead to 100% perfect R-wave recognition in the presence of high levels of noise. That is an unattainable goal.

What we are confident is an attainable goal, based on performance so far, is a level of accuracy that produces minimal distortion of the HRV metrics. That will be tested when data collected with the relevant environmental temperature extremes becomes available.

Time Performance of the Software. At this point in project, the first year, we are not yet at the stage of being extremely concerned about the computational burdens of the algorithms.

Nonetheless, we are mindful of the fourth objective of our proposal, "... to develop RT collection software of HRV in different modalities and RT computation of the rules using modest capability processors." The key word being modest capabilities. Our goal is that in the 6 months of the research program we will optimize the software and reduce all unnecessary overhead, so that it will be consistent with the data and processing capabilities of modern, rugged, very low-power, low-weight processors.

In this context, some comments on the current performance and test-bed are in order. The PI deliberately tests all the software on an relatively outdated 800 MHz P-III computer, rather that on the more contemporary 2.6 GHz P-IV desktop that will be used for A/D data acquisition for this project, or any faster workstation, lest that gave a false sense of performance of the software.

As an example, in that 800 MHz computer, the current version of ECG16a (as of the date of preparation of this report) can process a relatively "clean," noise-free file containing 480 minutes (8 hr) of recorded Lead II ECG and identify the 21,386 R-waves it contains using a Frequency bandpass filter as Filter 1, a Coiflet as Filter 2, and our Modified Pan-Tompkins algorithm for R-wave detection, in slightly over 1 minute, with ~ 99% accuracy. With the same filter settings, but using the Correlation algorithm for R-wave detection, the file can be processed in slightly over 2 minutes, with about ~ 98% accuracy.

The accuracy figures are available when the ECG16a program is run on data whose R-waves have been previously identified and vetted; i.e., that has been previously independently scored by an expert. These numbers, in turn, together with their impact on the HRV metrics that are computed by the ECG16a program, for part of Task 3.

For a number of technical reasons, it is not possible to extrapolate this performance to real-time R-wave identification and HRV metric computation, the next step which we are now investigating. Clearly, our R-wave identification is far faster than it needs to be to keep up to the fastest human heart rate.

Tests with data with noise also indicate that the algorithms so far are fairly robust in terms of where, if noise makes them "lose" where R-waves are in a signal, the algorithms can "auto-adjust" and "re-synch" without human operator intervention; in fact ECG16a does not need any information to auto-start and begin detection of R-waves.

We also know that for clean data, the computational cycles required to compute the time-domain HRV metrics we have chosen to implement in ECG16a and examine first (SDNN, RMSSD, NN50, pNN50) and the Fourier spectral HRV metrics (VLF, LF, HF, LF/HF, Total Power, %VLF, %LF, %HF) can be computed and updated while the ECG16a is also identifying R-waves. What we don't yet know are what improvements and choices of pre-filters (band-pass, Kaiser window) and wavelet/fractlet filter we will need when the data are very noisy and the software has to "re-synch" often. That is currently being investigated with noisy data sets.

TASK 3. Noise Sensitivity Analyses. As stated in the approved SOW, the strategy to be pursued in Task 3 was to use pre-existing ECG records, both short-term (4 to 10 min) and long-term (18-24 hr) records to examine candidate algorithms for imputation of HRV in the presence of noise, artifacts or missed-beats. The logic was that these candidate algorithms would subsequently be tested against real data under hypo- or hyperthermia and movement conditions with Task 4 data.

The core of Task 3 are the same two sets of "Gold Standard" ECG records that have been painstakingly scored and validated and used in Task 2 as well. As noted above in the description of Task 2, the first set is used to test, and the second to validate the algorithms.

Since so much hinges on the accuracy of the R-wave assignments in the records that comprise the "Gold Standards," a fixed procedure was implemented. Each record in a set was first computer-scored by an earlier version of our software using our optimized Pan-Hamilton-Tompkins R-wave detection algorithm. As our software improved in the course of the study, new versions were used to improve the efficiency of the process. Each computer-scored file was then examined by one staff member using tools built into the software, which flag in both tabular and graphical fashion, every minute where there is an R-R interval that is "suspect," or where the algorithm has labeled an R-wave assignment of less than high-confidence. Each and every one of those minutes was visually inspected by the staff member, and corrected if necessary. Finally, the file is re-checked by the PI before a file is moved or vetted into the "Gold Standard" set.

The first, test "Gold Standard" set consists of:

- Sixteen, 8-hr ECG records of volunteers sleeping (digitized at 256 Hz)
- Six, 8-hr ECG Holter records of fully-ambulatory volunteers (digitized at 180 Hz)
- Four, 45- to 90-min ECG Holter records of volunteers engaged in heavy exercise (digitized at 180 Hz)
- Fourteen, 20-min ECG records 10 min-supine, 10-min standing, continuous (digitized at 256 Hz)

The second, validation "Gold Standard" set consists of:

- Fifteen, 8-hr ECG records of volunteers sleeping (digitized at 256 Hz)
- Six, 8-hr ECG Holter records of fully-ambulatory volunteers (digitized at 180 Hz)
- Three, 45- to 90-min ECG Holter records of volunteers engaged in heavy exercise (digitized at 180 Hz)
- Fourteen, 20-min ECG records 10 min-supine, 10-min standing, continuous (digitized at 256 Hz)

Noise sensitivity analyses. In the approved SOW, the PI proposed to begin with ectopy-free and noise-free data, so that "true" HRV may be computed and errors and biases introduced by correction methods evaluated. The originally-proposed plan was as follows: Starting with the clean test ECG record, and its conventional HRV metrics, the PI would use carefully splice, at random positions in the record, segments of noisy data from other series in his database. This would create a synthetic new ECG record with large and unpredictable components of noisy records, but comparable HRV values except for the "noisy" segments. These synthetic records would then be used to test various algorithms to preserve the "true" HRV parameters in the presence of noise, and also to test algorithms for the faithful detection of R-waves in the presence of noise.

In fact, during the initial phases of performance of Task 3, it was found that a minor modification of our software enabled us to obtain some of the requisite data for the noise sensitivity analyses without the use of splicing. Virtually all of our records contain some ectopic beats and/or noisy segments. Even the cleanest of the records obtained in the laboratory overnight contain noise and movement-induced myopotentials, this is particularly true when the volunteers move around in their sleep and are in Stage IV or REM sleep; we know this because some of these previously-collected ECG records involved polysomnography and the EEG records had been scored for sleep stages.

Our subcontractor, Ulrich Associates, Inc., modified our software so that we could mark R-waves as ectopic by standard electrocardiographic criteria, and also in noisy segments we could mark "begin noisy-bad segment/end noisy-bad segment." When the R-waves and their time of occurrences were saved, in fact two sets of R-waves, and R-R intervals were saved separately. The first is the conventional file that contained all of the R-R intervals, whether the R-waves corresponded to normal R-waves, ectopic beats mis-identified as R-waves, or noisy segments. That file, when processed "blindly" by the HRV-metrics software, would produce HRV metrics that are "contaminated" by errors in R-wave assignments and noise (Xia et al., 1993; Lippman et al., 1994; Berntson and Stowell, 1998).

In the second file, the "Good R-R" file, the marks for ectopic beats and bad segments had caused the data to be handled in a specific way by the modified software. Upon encountering a begin ectopic/end ectopic mark, the last normal R-wave prior to the ectopic beat and the first normal R-wave after the ectopic beat become "identified" in time, so there is no artifactually long segment. This is the "deletion" method recommended by Lippman et al. (1994) as being the one that gave the least error in HRV metrics. To quote from their Methods section: "Deletion: ectopic beats were deleted from the R-R interval listing, and subsequent beats were shifted down in the sequence to take their place. Thus if there were initially N R-R intervals, of which K represented ectopic beats, then the resulting R-R interval list after deletion would contain N-K R-R intervals." (Lippman et al., 1994).

The segments that contained bad or noisy data were treated in the same way by the "deletion" method. In this fashion, the two files generated from the same data set could be compared with respect to their HRV metrics, against the number and severity of noise segments. We are currently completing these analyses for the test set of the "Gold Standard" files, and will proceed with more extensive tests based on the originally-proposed splicing strategy.

Surrogate data. It was indicated in the approved SOW that our candidate algorithms (developed in Task 2), and with known noise sensitivity parameters tested in Task 3, would be subsequently be tested against real data under hypo- or hyperthermia and movement conditions with Task 4 data. Under the projected time-line of the Tasks, the PI would have had some data from Study 4 that included ECG records with volunteers engaging in strenuous movements, stomping, etc., to challenge the recognition algorithms and help validate the noise analyses for the HRV metrics. As noted in the progress report for Task 1, approval of the Protocol has been delayed by the HSRRB, so as of this date no new data are available from Task 4.

For this reason, in order to maximize the progress while dealing with the unavoidable delay in Protocol approval, the PI focused attention in both Task 2 and Task 3 on several pre-existing ECG records in his database that had volunteers engaged in strenuous exercise, and which therefore had ECG traces contaminated with considerable myoelectric noise. These records have far more myoelectric noise than the 18-hour Holter records of ambulatory volunteers going about their daily activities, very few of whom engaged in strenuous activities at that time.

In the strenuous exercise records, the volunteers had engaged in the exercises that are part of their normal fitness program. One particularly useful set of records for algorithm development, validation, and noise sensitivity analyses were those recorded from a volunteer who practiced martial arts, in particular Tae-bo. In this discipline, there is a sequence of movements that involves standing, tensing up, thrusting, and stomping which introduce very different forms of noise into the ECG record, and which present very similar R-wave recognition challenges to those that will occur with the data that will be collected in Task 4. All of the aforementioned motions in Tae-bo are as were described to the PI who has not witnessed them personally, but who queried the volunteer as to the annotations that accompanied the Holter record.

The next series of figures illustrates the R-wave recognition challenges presented by the data in this file, which are representative of the challenges that have to be overcome successfully for implementation of HRV in RT in field-like conditions. One of the important points to keep in mind when considering those figures is that it is relatively simple to design filters that will recognize R-waves for individual types of noise. But all of the diverse examples presented arose from one type of activity, and need to be effectively handled by one algorithm and one filter in any kind of practical implementation.

All of the figures that follow are screen captures from ECG16a, analyzing a Tae-bo exercise record. As noted previously, in ECG16a, there are four traces. The first trace is the raw ECG. In the following series of figures, it will be important to notice the units in the raw trace, which although they have not been converted to microvolts of ECG signal, they are nonetheless exactly proportional to the signal recorded by the Holter monitor at all times. The software auto-scales the signal for ease of display purposes, and attention to the scale underscores the wide variation in amplitudes seen under exercise conditions, even on a second-to-second basis, something that is not encountered in routine laboratory conditions or even in stress tests that use predictable Bruce multistage treadmill protocols.

The second trace displays the output of the first filter, which in the present example is a low-pass filter with a steep (fourth-power) characteristic. The third trace has the output of the second filter, in this case a Coiflet-6 fractlet of level 2.5, as well as other information. On that third trace, the PI-identified R-waves appear as triangles (open or solid), while R-wave assignments that are made by the ECG16a software without using the information in the triangles appears as vertical solid bars; black bars as high confidence identifications, while red or green bars are indications of self-reports of problems with the auto-recognition algorithm. This gives an immediate, minute-to-minute reflection of places where the recognition algorithm failed or succeeded, in addition to tabular summaries for the entire file that can be invoked via menus. The fourth trace is the output of the integrator (for the Pan-Hamilton-Tompkins R-wave recognition, as is the case here), or the output of the correlation function (if the correlation R-wave recognition method has been chosen).

Figure 4 illustrates the baseline for this volunteer, immediately prior to any exercise, and showing the best signal-to-noise ratio that could be expected that particular day for the Holter monitor; Lead II ECG.

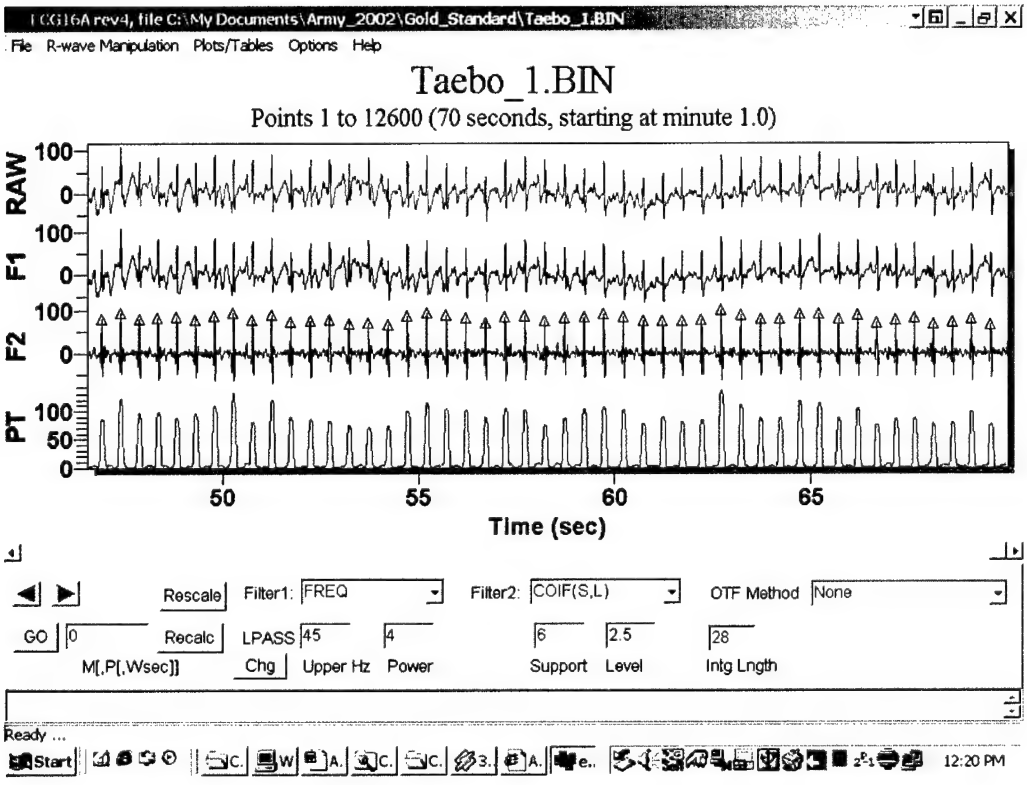

Figure 4. Lead II ECG of volunteer, baseline, immediately prior to any exercise.

Figure 5 shows a minute after Figure 4, where there is a tensing up of the muscles, and rhythmic myoelectric noise about 2X-3X greater in amplitude than the R-wave amplitude to be detected is introduced into the record; notice the change in scale in the raw signal, which has gone from +/- 100 to +/- 250. The frequency filter was not appropriately tuned for this particular event, but the fractlet clearly picks up the relevant R-waves.

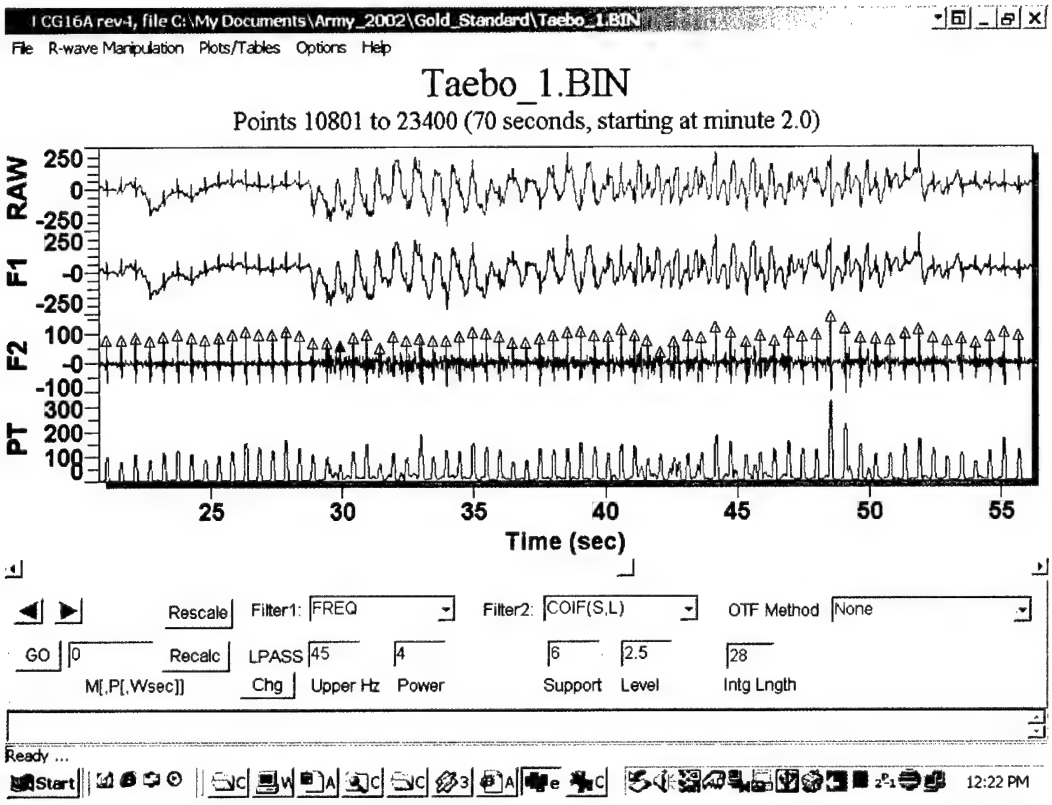

Figure 5. Lead II ECG of volunteer, brief muscle tensing up.

Figure 6 illustrates a high-amplitude, low-frequency oscillation that appeared in the record (not clear to the PI what aspect of the exercise elicited this noise). It is clear that changing the first filter from a low-pass to a band-pass filter can eliminate that kind of noise, if that is the only kind of noise in the record. However, that change also changes the spectral characteristics of the QRS complex to be recognized. What is reassuring in Figure 6 is that, as is clear from the sharp peaks in the second trace, the fractlet filter removes that kind of oscillation efficiently.

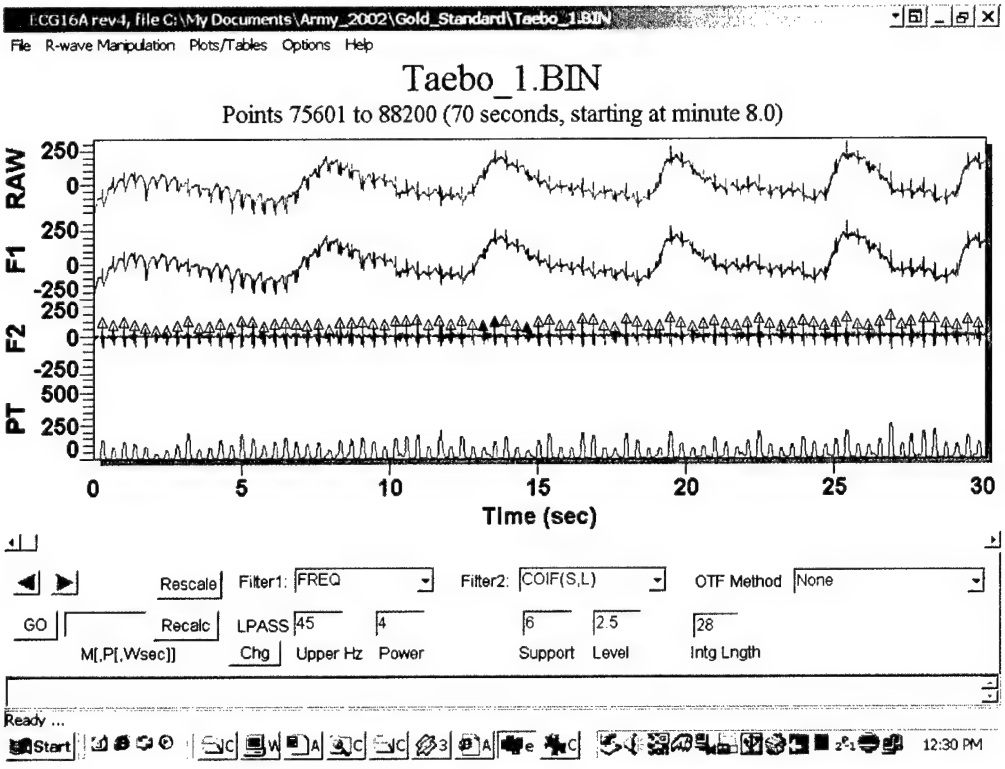

Figure 6. Illustration of a high-amplitude, low-frequency myoelectric oscillation superimposed on the ECG.

Figure 7 illustrates a very different type of noise that is induced in the ECG record when the volunteer engages in the vigorous "stomping" motion. The magnitude of the artifact is over 4X the amplitude of the R-wave. Again, the fractlet filter removes that kind of artifact efficiently.

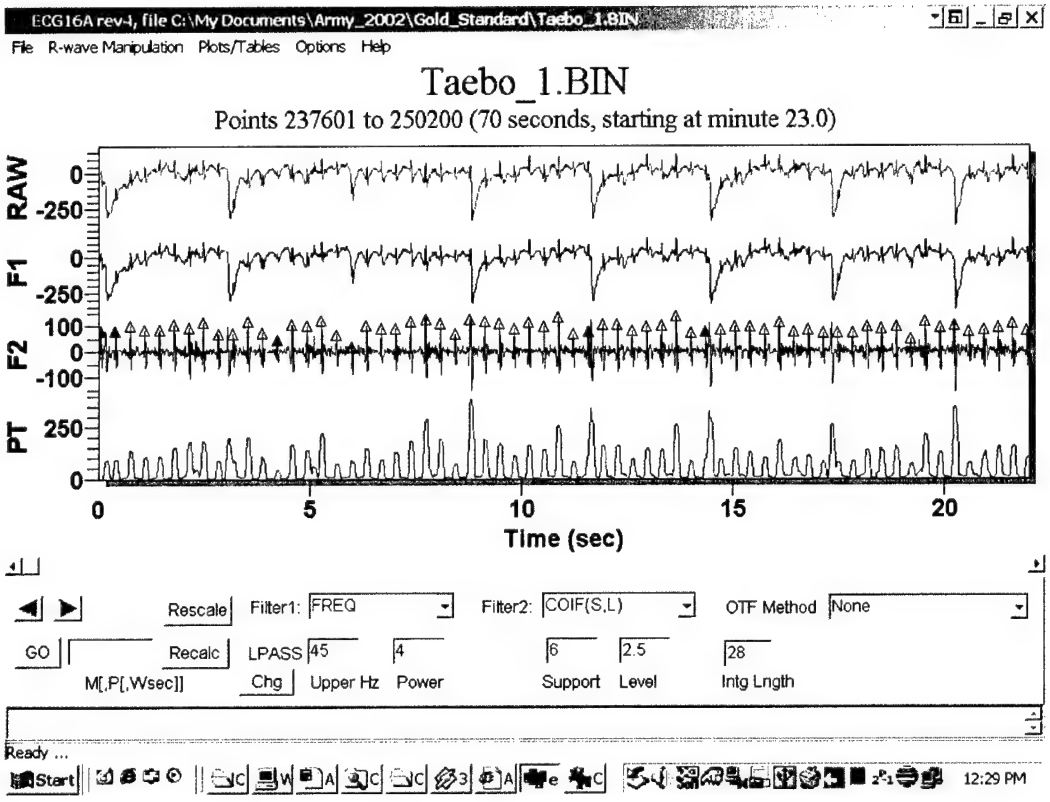

Figure 7. Illustration of an artifact superimposed on the ECG due to vigorous stomping.

So far, Figures 4-7 have just given visual indications of the challenges for the R-wave detection software, and visual indications of the efficacy of the fractlet filter. The next series of figures address the performance of the most current version of our software, ECG16a, on this file, using automatic recognition of the R-waves with no human intervention. In this mode, the software auto-initializes, and does not use the information used to display the triangles that are seen in the figures. The R-wave recognition algorithm's placement of the R-waves is shown by vertical bars. For the sake of brevity, one example will be shown pictorially, and the rest will be shown with tabular summaries.

Figure 8 illustrates the R-wave recognition algorithm's placement of the R-waves for the high-amplitude, low-frequency myoelectric oscillations superimposed on the ECG that were illustrated on Figure 6. As can be seen in that illustration, using our optimized, Modified Pan-Tompkins algorithm, the R-wave recognition is flawless.

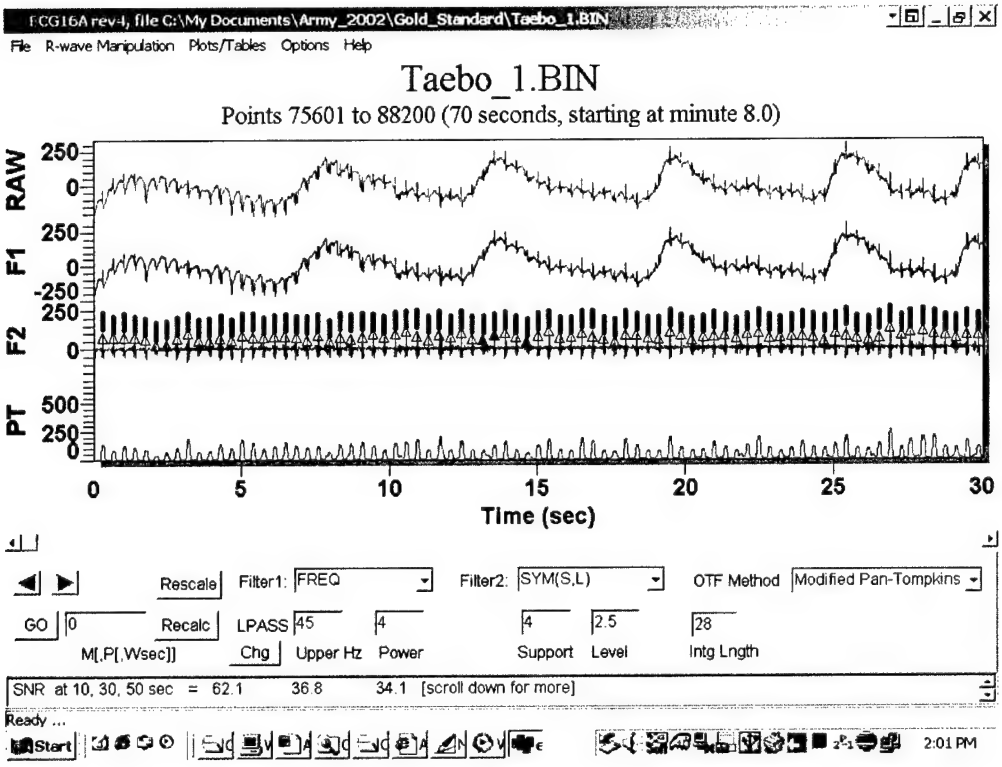

Figure 8. R-wave recognition by a Modified Pan-Tompkins algorithm. Same data segment with high-amplitude, low-frequency myoelectric oscillation superimposed on the ECG as illustrated in Figure 6.

The performance of ECG16a on the entire file can be best appreciated from the summary statistics generated, which are shown in Figure 9. The button on checked on top "On-the-fly" is laboratory jargon for the results of the R-wave recognition performed by the R-wave recognition algorithms (Modified Pan-Tompkins or Correlation with Filtered Data) working without human intervention. The summary indicates the file duration in minutes, and the number of "valid" R-waves, meaning those determined to be valid from the "Gold Standard" scoring reflected in the triangles on the screens. The number of "cold starts" is laboratory jargon for the software auto-starting R-wave recognition without human intervention. There will always be one "cold start," the beginning of the file, or when the software deals with RT data, the beginning of the data stream. A value of no greater than one means that at no time did the algorithms "lose" track of R-waves and had to "restart" detection. For example, a value of two would have indicated that, in addition to the beginning of the file, at some other point in the file the algorithm had to re-start, and another tab in the Rwave Summary screen would indicate at what minute that occurred.

The "number of false positives" and "number of false negatives" directly addresses the noise sensitivity analysis goals of Task 3. They are computed by comparing the R-wave detection algorithm (Modified Pan-Tompkins) assignment of R-waves, independent of any human input,

against the validated assignments for this file. As can be seen from that figure, and for this very challenging file, those numbers were 48 R-waves out of 7097 and 43 out of 7097, respectively. That yields a false positive error rate of 0.68%, and a false negative error rate of 0.61%. We maintain those are excellent numbers for such noisy data.

The Correlation with Filtered Data R-wave recognition algorithm does not perform as impressively. This method is of more recent vintage in our analytical arsenal, and has not received as much attention for optimization as the Modified Pan-Tompkins method has received from Prof. Davis, Mr. Ulrich and from the PI. For this file, while the false positive and false negative error rates were comparable to those of the Modified Pan-Tompkins algorithm (approximately 0.5%), the Correlation algorithm required 4 "cold starts" in the middle of the file, in addition to the start off the file. We do not yet know if the method is intrinsically inferior, whether it simply needs more work to improve performance, or whether an optimal detector can be crafted from an adaptive detector that switches between Pan-Tompkins and Correlation mode depending on the noise levels and characteristics of the data.

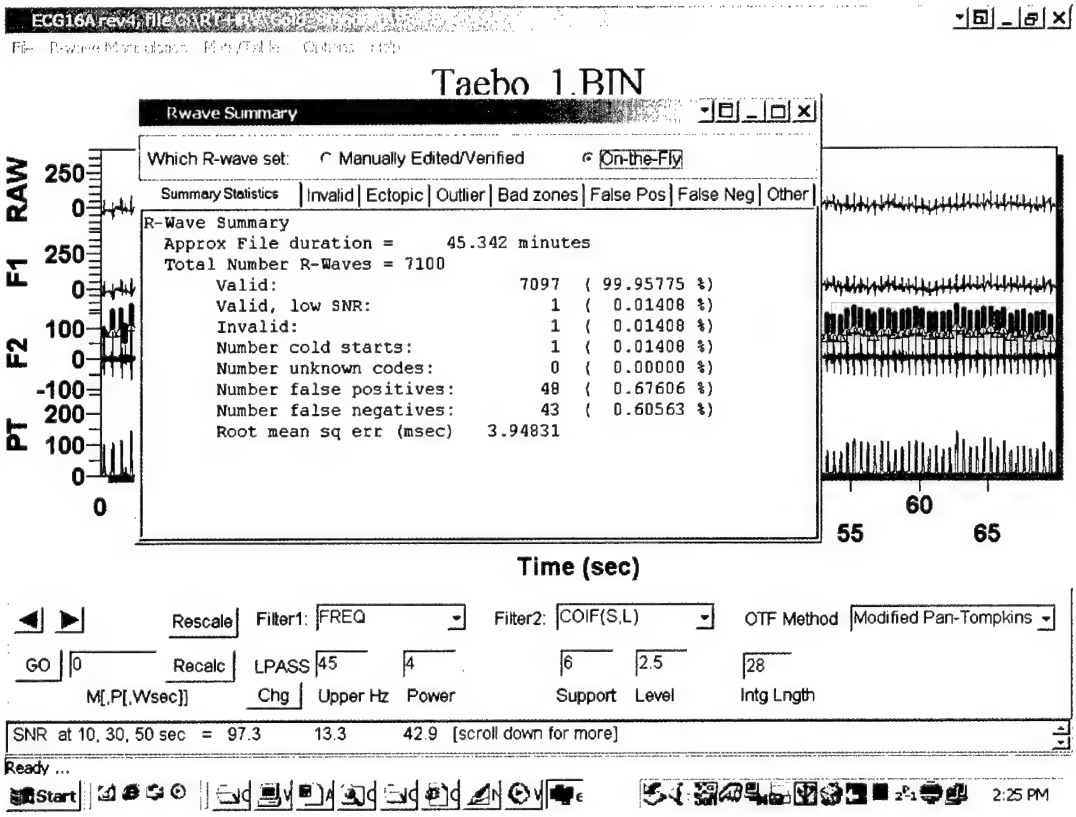

Figure 9. R-wave Summary Statistics.

Error rates and impacts on HRV metrics. This software also enables us to, with a simple click of a mouse, to generate the time-domain and spectral HRV metrics for either the human-

validated, "Gold Standard" triangles, or for the R-waves computed by the R-wave detection algorithm, either the Modified Pan-Tompkins or the Correlation with Filtered Data. Since by definition the "Gold Standard" is error-free, while the R-wave detection algorithms each will have, for each data set, a different false positive and false negative error rate, we can immediately see the impact of these error rates on the HRV metrics.

Figure 10 illustrates the first of several outputs provided by the program for the HRV metrics. This one is along the lines of a "sanity check" for the PI and the programmers to insure that the program is providing correct answers at all stages, as well as to provide the PI with full control, with visual feedback, of the data conditioning procedures that go into the metric computations. Since there is relatively little standardization of how investigators compute and report HRV metrics in the peer-reviewed literature, and one of the approved tasks in the SOW is to develop custom metrics, a great deal of flexibility has been built into the software.

The first item concerns basic units. The raw unit of measurement for all HRV is the R-R interval, in msec, and that is the unit used for time-domain metrics. But a problem immediately arises for Fourier spectral frequency metrics. Fourier transforms (like the Fast Fourier Transform, or FFT) are applied to time-series of data that are equally-sampled in time. A series of R-waves, or differences in onset of R-waves, are most definitely not equally sampled in time: R-waves happen whenever the heart beats. Therefore, it is customary to take the series of interbeat (R-R) intervals and re-sample them using one of several fixed moving window schemes to provide a time series that is equally-sampled. Some investigators then report the results as beats per second or beats per minute, while others report seconds-per-beat. This is not quite a trivial difference, since it there are statistical distributional differences between these two variables. We have incorporated the ability to use either method. In addition, some investigators report re-sampling at 0.5, 1 or 2 Hz. There are pros and cons to finer or coarser sampling, up to reaching the Nyquist limit of the signal, and these in turn relate to mean heart rate. We have implemented a method that allows the PI to re-sample the R-R intervals at 0.1 Hz intervals, and we will examine, for the range of heart rates that will be present in the environmental extremes that we will study, whether a clear optimum re-sampling rate exists. Please note that in Figure 10 below the screen mis-labels that as "Sampling rate." That, of course, is not the rate at which the ECG was sampled, and will be correctly re-labeled "Re-sampling rate" in a later revision of the software.

The next area of non-uniformity has to do with accumulation time for the metrics. While it is customary to use 5-minute accumulations for time-domain metrics, virtually every conceivable time period and method of analysis has been used for Fourier spectral frequency metrics. Some investigators will use segments as short as 1 or 2.5 minutes, others will use 5 or 10 minutes. Yet others will use the Welch procedure, where short FFTs are computed of partially-overlapping segments and averaged. We have implemented a variable accumulation time, so the PI can experiment with times as short as 1 minute or as long as the entire record for the development of custom metrics for the data obtained upon exposure to environmental extremes.

The screen illustrated in Figure 10 below also illustrates in graphical form the waveform being submitted to the FFT. The screen as it stands is redundant in one aspect, in that the current detrending algorithm also removes the mean. It is typical practice to always remove the mean

and linear trend, since otherwise there is artifactual leakage into the low-frequency bins. The signal is zero-padded to the next power of 2 (for computational FFT purposes), and a Hamming window is applied to avoid edge artifacts. The resulting one-sided power density from the FFT is displayed, with a summary of a typical computation of VLF, LF and HF bands, as well as "Total Power" on the right-hand side. It is emphasized that this is still the "sanity check" screen, and that more extensive information is available and will be shown below. It should be also noted that were are using the strict Consensus definition of Spectral Total Power (as implemented in the Task Force (1996) Report), which excludes any power below 0.003 Hz and above 0.4 Hz, and therefore this "Total Power" will always be less, as opposed to mathematically identical to SDNN.

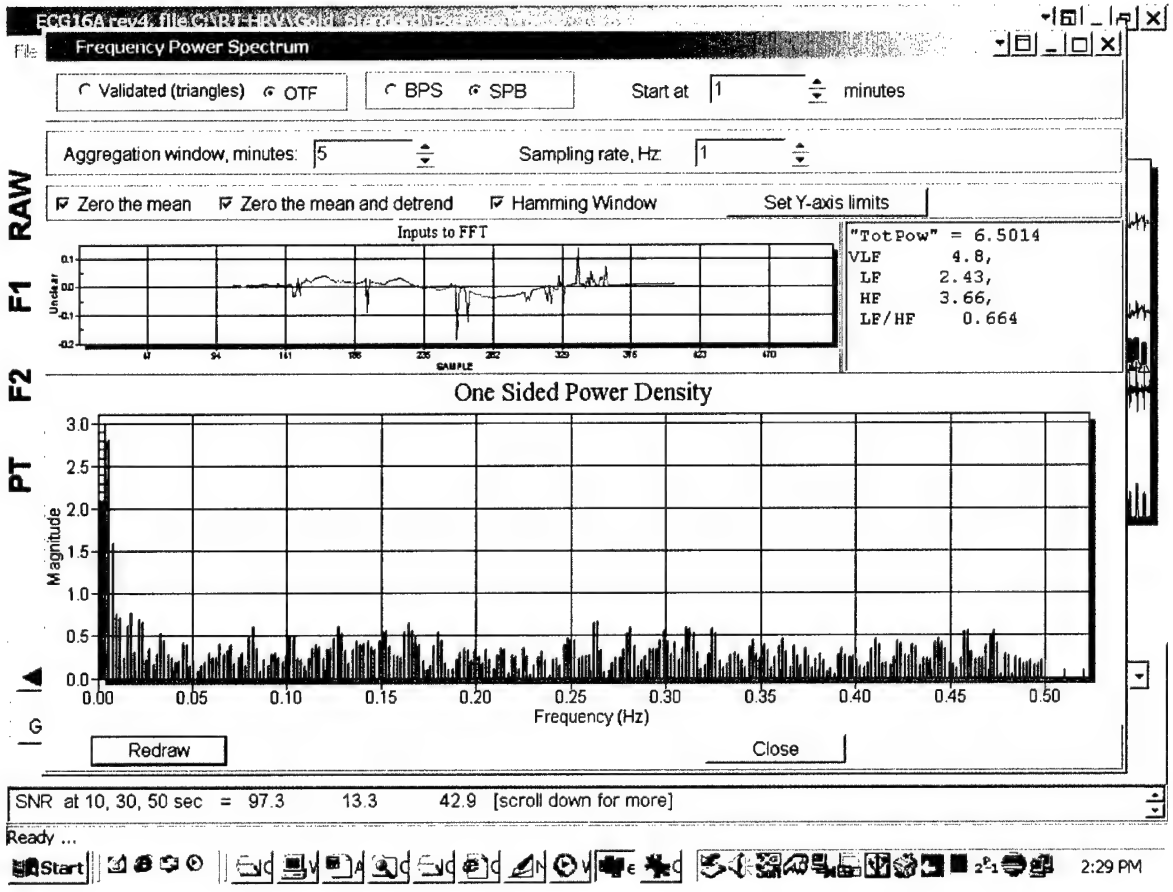

Figure 10. FFT Summary Screen.

Figure 11 shows one of several displays of the resulting time-domain and Fourier spectral frequency HRV metrics. The screen was designed to facilitate the PI's task of extracting information from the data, especially as the data from the volunteers exposed to environmental temperatures extremes became available, since so little is known about what to expect in terms of HRV changes.

As with previous panels, the PI has control over display of the "Gold Standard," validated data, the points identified by the R-wave detection algorithms, or in this case, both at once. The button on the upper right-hand corner permits the export of all of the numerical data (the metrics) to programs that accept ASCII input, like Excel and most statistics programs, like Systat and SAS. There is control of which of the time-domain and Fourier spectral frequency metrics to be displayed, in mix-and-match fashion. While in theory all boxes could be checked, in practice more than 5 or 6 in one plot becomes cumbersome, but the flexibility is welcome. Where appropriate (like for LF, HF), there is the choice to express the quantities as absolute or as percentages normalized against total power, since sometimes physiological patterns are apparent only upon those normalizations.

The graphic that occupies the bottom half of the screen is self-documenting with respect to the source file. It gives the full path information to its location either in the local hard disk or on the network, to facilitate matching a screen capture with numerical data and reduce possible errors in both data-exploration and data-analysis phases. Some of the various axis options will become clearer by comparing Figure 11 and Figure 12. In Figure 11 the variables (Frequency-domain LF, HF and Total Power) and time-domain pNN50 are plotted on a common Y-axis.

One aspect that immediately becomes interesting in this record, from a scientific and HRV metrics point of view, is that there are 2 very noteworthy segments, with one "classic" segment serving as a comparison. The segments to be noted are between minutes 11 and 16, 16 and 20, and 35 to 40. The "classic" segment is that between minutes 11 and 16, where it can be seen that the increase in Total Power is perfectly tracked, and could be quantitatively accounted for, by LF and HF.

In marked contrast, immediately after minute 16, between minutes 16 and 20, there is a discrepancy. There is another rise in Total Power, that is not accompanied by changes in LF or HF; both these quantities remain flat and low. Since the Consensus (Task Force, 1996) TotalPower is defined as between 0.003 Hz and 0.4 Hz, LF is defined between 0.04 and 0.15 Hz and HF is defined between 0.15 to 0.40 Hz, it is tempting to dismiss the discrepancy as simply missing the VLF power (between 0.003 and 0.04 Hz) which is not plotted in this illustration. The same argument could be made for the discrepancy between minutes 35 and 40. As will be discussed below, that simple explanation is scientifically not tenable.

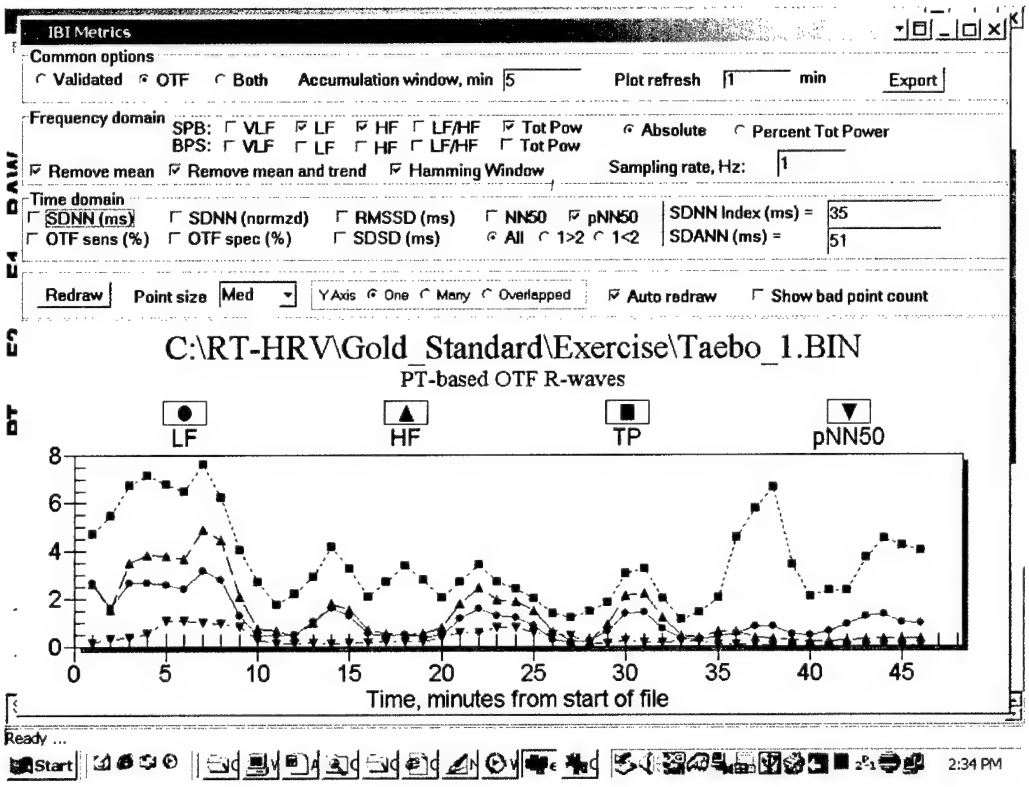

Figure 11. HRV Metrics Screen, Single-Axis.

Figure 12 shows another of the several displays of the resulting time-domain and Fourier spectral frequency HRV metrics from these data, with extra information to help dissect the issues surrounding the discrepancies in the segments between minutes 16 and 20 and minutes 35 and 40. The multi-axis feature of the plot has been activated, to be able to visualize, in addition to the previously-plotted metrics, also the time-domain SDNN (which in single-axis mode, due to its units, tends to make it difficult to see the other axes). The reason SDNN is important in helping sort out the apparent discrepancy, in addition to a direct plot of VLF (which does not account for the discrepancy in this file), is a more general property. As previously noted, if the spectral data are expressed in the proper, time-based units upon re-sampling, it is a mathematical theorem that Total Power over the entire Nyquist interval and SDNN are mathematically identical. However, the Consensus document (Task Force, 1996) recommended defining Total Power not from 0 to 0.5 Hz (the Nyquist interval), but from 0.003 Hz to 0.4 Hz. So, it is not only the case that Total Power is always less than SDNN, but the sum of the Consensus-defined bands, ULF (not relevant for our studies, since our data collection is not circadian), VLF, LF, and HF still leaves the 0.4-0.5 Hz band unaccounted for.

As can be seen in Figure 12, in both of the segments in question, between minutes 16 and 20 and minutes 35 and 40, SDNN and Total Power fail to track. This illustrates that, as has been seen in more recent applications of HRV, the original Consensus (Task Force, 1996) bands like

LF and HF were defined with the best available basic science and clinical knowledge, but some applications, like detection of apnea, require definition of new bands that have frequency ranges outside those envisioned in the Consensus document. Whether that turns out to be the case for the thermal-derived endpoints of interest in this study remains to be determined empirically. Be that as it may, Figure 12 illustrates that the project has the software with the flexibility to explore these questions.

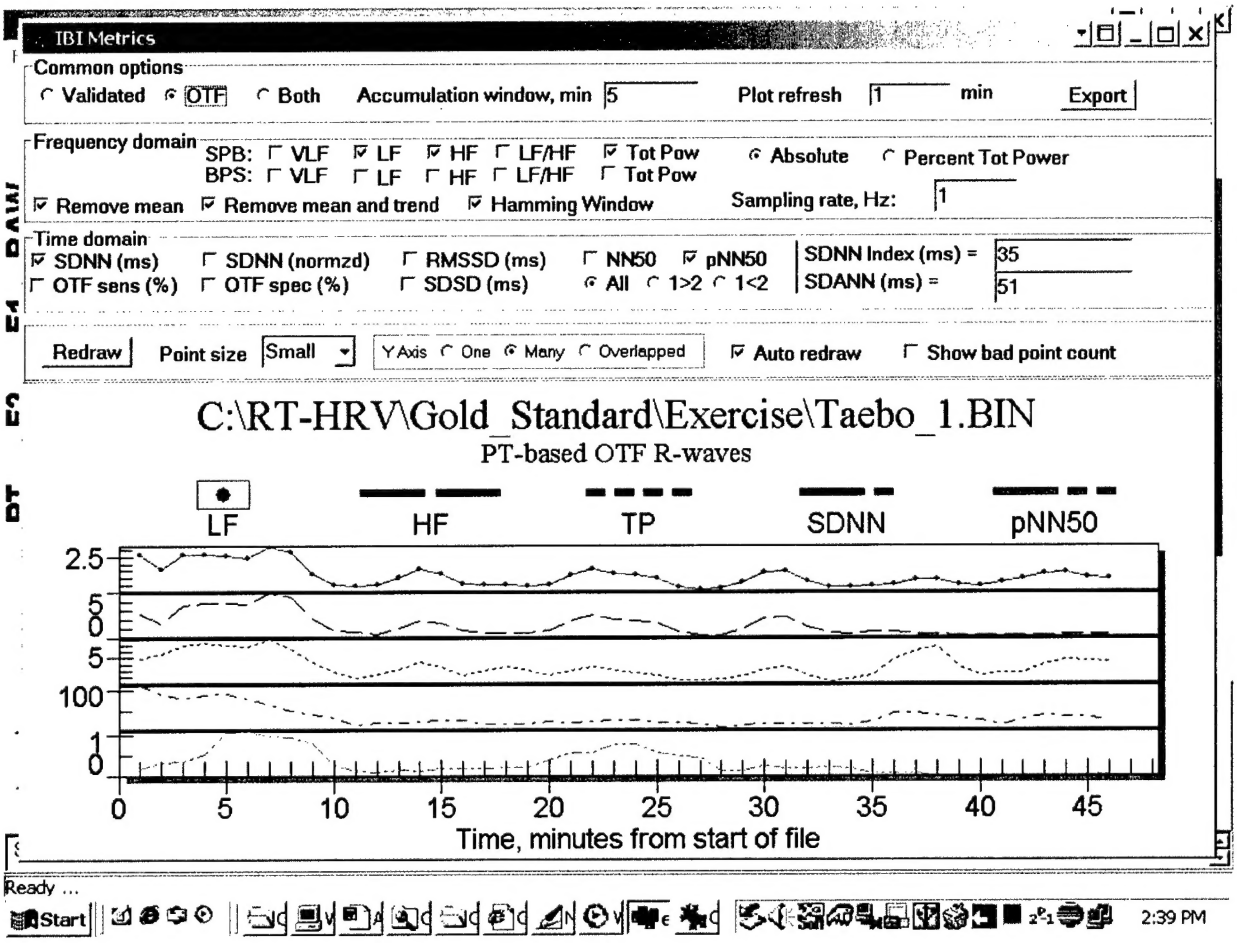

Figure 12. HRV Metrics Screen, Multi-Axes.

TASK 4. Performance of Human Studies.

Clearly, since we are awaiting approval from the HSRRB on our Protocol, we have limited progress to report on Task 4. However, work that can be completed prior to receiving that approval has been either completed or is well under way.

For example, the peer reviewers of the proposal had recommended that when the PI record data in the human subjects part of the study, he initially over-sample the ECG data in the

frequency domain and later use decimation if necessary. The logic behind this recommendation was that if data are acquired at, for example, 256 Hz sampling, then the Nyquist limit for that sampling, 128 Hz, is the highest information content that will ever be available. If, instead, the signal is originally sampled at 2,048 Hz, and later analyses indicate that frequencies in the 200 or even 500 Hz are of importance, they will be present and can be recovered, since the Nyquist limit for that recording would be 1,024 Hz. It has happened in the past decade in cardiology that new developments call for higher-resolution data, and older databases collected at slower digitization rates are inadequate. A case in point is the clinically-important and predictive phenomenon of T-wave alternans, whose accurate detection requires both a high-gain and sampling rates in excess of 400 Hz.

The PI took this advice to heart, and under Task 4 he and his staff have modified their data-collection software to accept variable digitization rates, from the past laboratory default of 256 Hz up to 20,000 Hz while simultaneously sampling, digitizing and writing to disk the 5 channels that will be required in this study. This software has been tested and has performed flawlessly at 20,000 Hz for data collection periods of up to 30 minutes with staff under exempt conditions. We are ready to record and store data as soon as approval is granted.

KEY RESEARCH ACCOMPLISHMENTS:

- Preparation and submission of Protocol for performance of study involving human volunteers that will examine cardiovascular and HRV responses to exposure to environmental temperature extremes.
- Development of a highly-sophisticated software suite that incorporates a number of features:
    - refinements in the Pan-Hamilton-Tompkins R-wave detection algorithm
    - blending with different pre-filters optimized for performance with the Pan-Hamilton-Tompkins algorithm
    - a new R-wave detection algorithm based on Correlation with Filtered Data
    - a wavelet library cross-validated against the commercially-available S-Plus library
    - a three-stage filtering testbed that allows visual experimentation
    - auto-start, auto-restart detection of R-waves without human operator intervention
    - relative insensitivity of R-wave detection process to myoelectric and movement-induced noise that is several-fold greater than the QRS signal itself
    - time-domain and Fourier spectral frequency HRV metrics
- Development of a fully-validated database of short-term and long-term ECG records for algorithm development and quantification of sensitivity of metrics with respect to noise and missed beats.

- Discovery, development and implementation of a new analytical tool (fractional order wavelets, or "fractlets") with excellent practical properties for identification of R-waves in noisy records.

REPORTABLE OUTCOMES

1. ECG16a software with the features outlined in the report and in the Key Research Accomplishments.

2. An interactive demonstration of the capabilities of the ECG16a software will be presented at the Technology Showcase of the Peer Reviewed Medical Research Program (PRMRP) Military Health Research Forum on April 28, 2004.

CONCLUSIONS

In conclusion, the progress to date in the project has been primarily in the area of software development, due to delays in approval of the protocol to perform the study with human volunteers.

In the area of software development, the progress has been excellent. We have accomplished all of the goals that had been set for this time period. The current version of our software is able to identify R-waves, without human intervention, in ECG records that are severely contaminated with myoelectric noise many times in amplitude greater than the amplitude of the QRS complex. The accuracy of identification is very good, with false-positive and false-negative error rates in the noisiest records tested to date of 1% or less. That same software is able to compute all of the standard time-domain and Fourier spectral frequency HRV metrics and display them in a flexible and useful way for data-mining and for development of custom metrics once data from the human studies with environmental temperature extremes becomes available.

The overall importance of our findings to date is that it is realistic to expect that, by the completion of this project, we will be able to deliver:

- An implementation of HRV that can run in RT
- Novel insights into the human responses to environmental temperature extremes as reflected in HRV, and
- New HRV metrics that may predict the approach to hypo- or hyper-thermia.

The importance of above points, for both military and civilian medicine, are as follows:

Recently, various branches of the U.S. Armed forces have begun to invest in the development of technologies for real time RT HRV. RT HRV is being explored as a way to provide field commanders with continuous information on the battle readiness and health status of troops. This is important for operational purposes and to best triage and deploy medical and other rescue units, in an effort to maximize the impact of the trauma response "golden hour" that in turn maximizes survival.

The "**golden hour**" refers to the well-known observation, in military and civilian trauma medicine, that the ability to provide proper medical attention during the first hour post-trauma (or onset of conditions like hypothermia or hyperthermia) has a disproportionately large, positive effect in reducing trauma-related morbidity and mortality. Basically, if it were possible to develop a small, lightweight, reliable device that would continuously monitor a warfighter's HRV and send a small amount of relevant information (either continuously or on demand) to a monitoring and command center, field commanders would be able to determine which of their warfighters needed immediate medical attention or evacuation from the war theatre.

Practical implementations of RT HRV currently face a number of hurdles. Some of these hurdles are obvious computational and engineering issues, and this progress report has outlined some of our progress in overcoming some of these obstacles. But equally important are physiological issues that reflect our incomplete understanding of which parameters of HRV, computable in RT, are actually useful in a trauma setting and predictive of subsequent detrimental physiological responses with respect to environmental temperature extremes. It is for these reasons that we will focus on the development of custom HRV metrics geared specifically to the prediction of hypo- or hyper-thermia, and to their validation.

REFERENCES:

Berntson, G. G., Stowell, J. R. ECG artifacts and heart period variability: Don't miss a beat! Psychophysiology 35:127–132 (1998).

Fleisher L.A., Frank S.M., Sessler D.I., Cheng C., Matsukawa T., Vannier C.A. Thermoregulation and heart rate variability. Clinical Science 90:97-103 (1996)

Lippman, N., K. M. Stein, B.B. Lerman. Comparison of methods for removal of ectopy in measurement of heart rate variability. American. J. Physiology. 267: H411-418 (1994)

Task Force. Heart rate variability: Standards of measurement, physiological interpretation and clinical use. Special report, Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Circulation 93:1043-1065 (1996)

Xia, R., Odemuyiwa, O., Gill, J., Malik, M., Camm, A. J. Influence of recognition errors of computerised analysis of 24-hour electrocardiograms on the measurement of spectral components of heart rate variability. International J. Biomedical Computing, 32, 223–235 (1993).